

GGCAGAACGGGCATAGTGAACGGACCGTGTGGTATGCTCCGCGATGTCTGAGGCGACTTGGTG  
GTACCGAGGAGGGACTTCAAAACATGACCTGCATTACAGAAGGGAAGCGGAGGTTAACACCACA  
CTCGAGGAGTTGTTACTCTATTTTATTTTCTTAATAAATCTATGCATATTGACTTTTGGGATGG  
TCAATCCACATATGTACTATTTAAACAAAGTTATGTCTGCTCTGTTTGTGGACACTTCTCTACC  
TGATGATGAAAGAAGCAGCTTTAGGTCCATTCGGAGCATAACTGAGTTTTGGAAGTTCATGGAA  
GGACCCCTCATCGACGGCTTGTACTGGGACTCGTGGTATGGCACCAAACAGCTGTACAGTGTGA  
AGAACAGCAGCCGCATCTACTACGAGAACGTTCTTCTCGGCATCCCCAGAGTGCGGCAACTGCG  
AGTCCGAAACAACACTTGCAAGGTCTACCCAGCTTTCAGTCCCTGGTCAGCGACTGCTACAGC  
AAGTACACAGTGGAAAACGAAGACTTCTCTGATTTTGGCCTCAAACGCAATCCAGAATGGACGC  
ACACGCCTTCTTCCCGCACTGCCCCATGGCACTGGGGGTTTGTGGCGTATACCGAGATGGAGG  
ATATATAGTCAAGTTATCAAAATCAAAATCTGAAACCAAAGCCAAATTTGTTGACCTTCGACTG  
AACAACCTGGATTAGCAGAGGCACCAGGGCTGTTTTTATTGATTTCTCCCTGTACAATGCTAATG  
TCAACCTGTTTTGTCATCATCAGGCTGCTGGCAGAGTTCCCTGCGACGGGTGGGCTCCTCACCTC  
CTGGCAGTTCTACTCTGTGAAGCTCCTCAGATACGTCTCCTACTACGACTACTTCATTGCCTCC  
TGTGAAGTCATATTTTGTATTTTCTCTTTGTCTTCATAATACAAGAACTGAGGAAAGTGAACG  
AGTTTAACTCTGCCTATTTTCAAGAGTGTCTGGAACCTGGCTGGAGATGCTGCTCCTGCTGCTCTG  
TTTTCTCGCCGTGTCTTTCTATGCATACTGTAACATGCAGAGCTTTCTCTTGCTTGGACAGCTG  
CTGAAAAACACTGACAGCTATCCCGACTTTTACTTCCTTGCATACTGGCACATTTACTATAACA  
ACGTAATTGCTATCACTATCTTCTTTGCATGGATAAAGATATTCAAGTTCATAAGCTTCAATGA  
GACAATGTCGCAGCTGTCATCAACACTCTCCCGCTGCATGAAGGACATCGTGGGGTTTCGCCATC  
ATGTTCTTCATCATCTTCTCTGCTTATGCCCAGTTGGGATTTCTGGTTTTTGGGTCACAGGTTG  
ATGATTTTTCAACTTTTCAAAATTCATATTTGCACAATTTTGAATTGTCCTCGGGGACTTTAA  
CTTTGCTGGCATCCAGCAGGCCAACTGGATCTTGGGGCCCATCTACTTCATCACGTTTCATCTTC  
TTTGTGTTCTTTGTGCTCCTGAACATGTTCTTGGCAATAATTAATGACACCTATTCTGAAGTTA  
AGGCTGATTATTCAATAGGCAGAAGACCAGATTTTGAACCTGGTAAAATAATTCAAAGAGTTG  
CTTTAATGTTCTCGAGAACTCAGACTCAAGAAAGCTCAAGCTAAAGAAGAAAAGAAAATGCAA  
ACCACTGACTTGGCCAGAGAGCCAGAAGAGAAGGCTTTGATGAAAGTGAGATCCAAGAGGCAG  
AGCAGATGAAAAGATGGAAGGAAAGGCTTGAAAAAAGTATTATTCTACAGAAATTCAAGACGA  
TTATCAGCCTGTCACTCAGCAAGAATTCCGAGAGCTCTTTTTATACGCGGTGGAGCTTGAGAAG  
GAATTACACTATGTCAGTTTAAACTGAACCAACTGATGAGAAAGCTGCACTAGCAGGCTGACA  
AGTGGAGTCATTTTATAAGAAAAGGCAACCGAAGAATTTTATTTCAGTATGAAGATTTTTTCCTC  
TCAATTTTCTTCAGTAATGCAAAAGAAAAACCAAAAGTAACCAGAAGTGCTTTTTATTTCAAA  
GTTCTTGAAGTAAAGAGTAAACTCTTGTCCTTTGCTAACAGCCGTGTCTGCAGTAAACAAT  
GAAGGAGCCTGCGTGTTCCTAAGTGTGGAGAGGATCTGCGGGAATGTGGAACAGCTTTCCTTG  
CCTACTGGAACCAACAAACAAGCACACAATGGGACTCTCTGAGTGCCTGACAAAGTGAACGCAAG  
TACAGCCAAGCACACATGGTGAACCTGTCAGGGAACACAAGCACTTTATGGCGTCACTTTCAAG  
GAACATATTTTATATGGATTTTGAAGAGTCTTGTGTTGCTGATAAGAACTTCAAGAAGTCTAAGC  
TTGGCTTTGATTCTCTTGATTCTTATATTCTCAAGCACCGGAACACGATCCTCCTTCTGGG  
CATTCTAGGGAAGATAAACTCTGTAAAGCAAAAAAAGAAAAAAGAAAAA (SEQ ID  
NO: 1)

FIGURE 1

MSEATWWYRGGTSKHDHLHYRREAENVNTTLEELLYFIFLINLCILTFGMVNPHEMYLKNKMSSL  
FVDTSLPDDERSSSFRSIRSITEFWKFMEGPLIDGLYWDWYGTKQLYSVKNSSRIYYENVLLGI  
PRVRQLRVRNNTCKVYPAFQSLVSDCYSKYTVENEDFSDFGLKRNPEWTHTPSSRTAPWHWGFV  
GVYRDGGYIVTLSKSKSETKAKFVDLRLNNWISRGTRAVFIDFSLYNANVNLFCIRLLAEFPA  
TGGLTTSWQFYSVKLLRYVSYYDYFIASCEVIFCIFLVFIIQELRKVNEFKSAYFRSVWNWLE  
MLLLLCFLAVSFYAYCNMQSFLLLGQLLKNTDSYPDFYFLAYWHIYYNNVIAITIFFAWIKIF  
KFISFNETMSQLSSTLSRCMKDIVGFAIMFFIIFSAYAQLGFLVFGSQVDDFSTFQNSIFAQFR  
IVLGDFNFAGIQQANWILGPIYFITFIFFVFFVLLNMFLAIINDTYSEVKADYSIGRRPDFELG  
KIIQSCFNVLEKLRLKKAQAKEEKKMQTTDLAQRARREGFDESEIQEAEQMKRWKERLEKKYY  
STEIQDDYQPVTQQEFRELFYAVELEKELHYVSLKLNQLMRKLH (SEQ ID NO:2)

FIGURE 2

underlined = deleted in targeting construct

**BOLD** = sequence flanking Neo insert in targeting construct

GGCAGAACGGGCATAGTGGAAACGGACCGTGTGGTATGCTCCGCGATGTCTGAGGCGACTT  
GGTGGTACCGAGGAGGGACTTCAAAACATGACCTGCATTACAGAAGGGAAGCGGAGGTTA  
ACACCACACTCGAGGAGTTGTTACTCTATTTTATTTTCTTAATAAATCTATGCATATTGA  
CTTTTGGGATGGTCAATCCACATATGTACTATTTAAACAAAGTTATGTCTCTCTGTTTG  
TGGACACTTCTCTACCTGATGATGAAAGAAGCAGCTTTAGGTCCATTCCGAGCATAACTG  
AGTTTTGGAAGTTCATGGAAGGACCCCTCATCGACGGCTTGTACTGGGACTCGTGGTATG  
**GCACCAACAGCTGTACAGTGTGAAGAACAGCAGCCGCATCTACTACGAGAACGTTCTTC**  
**TCGGCATCCCCAGAGTGCGGCAACTGCGAGTCCGAAACAACACTTGCAAGGTCTACCCAG**  
**CTTTCCAGTCCCTGGTCAGCGACTGCTACAGCAAGTACACAGTGGAAAACGAAGACTTCT**  
**CTGATTTTGGCCTCAAACGCAATCCAGAATGGACGCACACGCCTTCTTCCGCACTGCC**  
CATGGCACTGGGGGTTTGTGGCGTATACCGAGATGGAGGATATATAGTCACGTTATCAA  
AATCAAAATCTGAAACCAAAGCCAAATTTGTTGACCTTCGACTGAACAACCTGGATTAGCA  
GAGGCACCAGGGCTGTTTTTATTGATTTCTCCCTGTACAATGCTAATGTCAACCTGTTTT  
GCATCATCAGGCTGCTGGCAGAGTTCCTGCGACGGGTGGGCTCCTCACCTCCTGGCAGT  
TCTACTCTGTGAAGCTCCTCAGATACGTCTCCTACTACGACTACTTCATTGCCTCCTGTG  
AAGTCATATTTTGTATTTTCTCTTTGTCTTCATAATACAAGAACTGAGGAAAGTGAACG  
AGTTTAAGTCTGCCTATTTTCAGAAAGTGTCTGGAAGTGGCTGGAGATGCTGCTCCTGCTGC  
TCTGTTTTCTCGCCGTGTCTTTCTATGCATACTGTAACATGCAGAGCTTTCTCTTGCTTG  
GACAGCTGCTGAAAAACACTGACAGCTATCCCGACTTTTACTTCCTTGCATACTGGCACA  
TTTACTATAACAACGTAATTECTATCACTATCTTCTTTGCATGGATAAAGATATTCAAGT  
TCATAAGCTTCAATGAGACATGTGCGAGCTGTCAACACTCTCCCGCTGCATGAAGG  
ACATCGTGGGGTTCGCCATCATGTTCTTCATCATCTTCTCTGCTTATGCCAGTTGGGAT  
TTCTGGTTTTTGGGTCAAGGTTGATGATTTTCAACTTTTCAAAATTCATATTTGCAC  
AATTCGAATTGTCCTCGGGGACTTTAACTTTGCTGGCATCCAGCAGGCCAACTGGATCT  
TGGGGCCCATCTACTTATCAGCTTCTCTCTTTGTGTTCTTTGTGCTCCTGAACATGT  
TCTTGGCAATAATTAATGACACCTATTCTGAAGTTAAGGCTGATTATTCAATAGGCAGAA  
GACCAGATTTTGAAGTTGGTAAAATAATTCAAAAGAGTTGCTTTAATGTTCTCGAGAAAC  
TCAGACTCAAGAAAGCTCAAGCTAAAGAAGAAAAGAAAATGCAAACCACTGACTTGGCCC  
AGAGAGCCAGAAGAGAAGGCTTTGATGAAAGTGAGATCCAAGAGGCAGAGCAGATGAAAA  
GATGGAAGGAAAGGCTTGAAAAAAGTATTATTCTACAGAAATTCAAGACGATTATCAGC  
CTGTCACTCAGCAAGAATTCGAGAGCTCTTTTTATACGCGGTGGAGCTTGAGAAGGAAT  
TACACTATGTGAGTTTAAACTGAACCAACTGATGAGAAAGCTGCACTAGCAGGCTGACA  
AGTGGAGTCATTTTATAAGAAAAGGCAACCGAAGAATTCATTCAATGATGAAGATTTTTT  
CCTCTCAATTTTCTTCAGTAATGCAAAAGAAAAACCAAAAGTAACCAGAAGTGCTTTTT  
ATTTCAAAATTTCTTGAAGTAAAGAGTAAACTCTTGTCTTTGCTAACAGCCGTGCTG  
CAGTAAACCAATGAAGGAGCCTGCGTGTTCCTAAGTGTGGAGAGGATCTGCGGAATGT  
GGAACACTTTCTTGCCTACTGGAACCAACCAAGCACACAATGGGACTCTCTGAGTG  
CCTGACAAAGTGAACGCAAGTACAGCCAAGCACACATGGTGAACTGTCAGGGAACACAAG  
CACTTTATGGCGTCAACTTTCAAGGAACATATTTTATATGGATTTTGAAGAGTCTTGTTT  
GCTGATAAGAACTTCAAGAAGTCTAAGCTTGGCTTTGATTCTCTTGATTCCTTATATTC  
CTCAAGCACCGGAACACGATCTCCTTCTGGGCATTCCTAGGGAAGATAAAACTCTGTAA  
AGCAAAAAAAAAAGAAAAAAAAAAAAAAAA

FIGURE 3

## Gene Sequence Structure

\*

211 bp

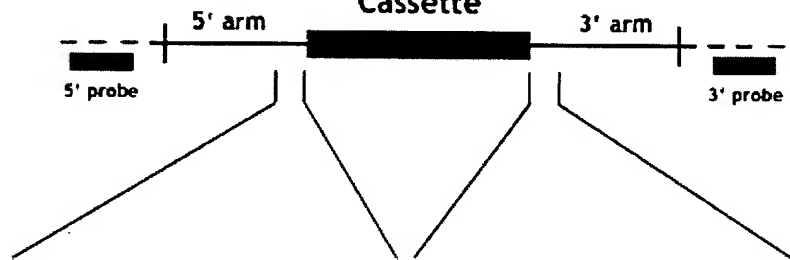
Sequence Deleted

355 bp

Size of full-length  
cDNA: 2486 bp

Targeting Vector\*  
(genomic sequence)

LacZ-Neo  
Cassette



Arm Length:  
5': 2.5 kb  
3': 1 kb

———— Targeting Vector  
----- Endogenous Locus

\* Not drawn to scale

5' >CATGCATACATGGGTGCACGT  
GCGTGCGTGCAAACACCCACACAC  
AGATAGATAAATGTAAAAAATAC  
TTAGGCAATGCTGTTTTTTTAAAG  
CTTGGGAAATTAATATTTTTGAAC  
ACTGCTGATGTGGGATTTTAAGTT  
ACAAATGTTAACTTGTTCTCCTT  
AGTGACTTTTGGGATGGTCAATCC  
ACATATGTACT<3'  
(SEQ ID NO:3)

5' >GTATGGCAACAAACAGCTGTA  
CAGTGTGAAGAACAGCAGCCGCAT  
CTACTACGAGAACGTTCTTCTCGG  
CATCCCCAGAGTGCGGCAACTGCG  
AGTCCGAAACAACACTTGCAAGGT  
CTACCCAGCTTTCAGTCCCTGGT  
CAGCGACTGCTACAGCAAGTACAC  
AGTGGAACGAAGACTTCTCTGA  
TTTTGGCCTCA<3'  
(SEQ ID NO:4)

FIGURE 4